DOCUMENT RESUME

ED 473 073 HE 035 664

| | |
|---|---|
| AUTHOR | Stoering, Juliette M.; Lu, Lina |
| TITLE | Combining the National Survey of Student Engagement with Student Portfolio Assessment. AIR 2002 Forum Paper. |
| PUB DATE | 2002-06-00 |
| NOTE | 15p.; Paper presented at the Annual Meeting of the Association for Institutional Research (42nd, Toronto, Ontario, Canada, June 2-5, 2002). |
| PUB TYPE | Reports - Research (143) -- Speeches/Meeting Papers (150) |
| EDRS PRICE | EDRS Price MF01/PC01 Plus Postage. |
| DESCRIPTORS | *Accountability; *College Students; General Education; Higher Education; National Surveys; Portfolio Assessment; *Portfolios (Background Materials); *Scoring Rubrics; Urban Universities |
| IDENTIFIERS | *National Survey of Student Engagement; Student Engagement |

ABSTRACT

Two assessment activities have been particularly useful at one urban university with an innovative general education curriculum: the National Survey of Student Engagement (NSSE) and the use of rubrics to score student portfolios. This study analyzed NSSE data in relation to the portfolio scores on four general education learning goals in an attempt to cross validate self-report and authentic learning assessments. The project was a collaboration between faculty in the general education program and the institutional research office. The sample consisted of 32 students who had completed portfolios as a requirement of their freshman year general education program. The learning goal rubrics applied to the portfolios and items on the NSSE appeared to have similar content, but findings show little correlation between the two sets of data. Despite the lack of significant results, the outcomes were valuable. Identifying the connections between diverse data sources and analyzing them in ways that respond to demands for accountability elevates the role of institutional researcher to that of partner in the university's attempts to provide authentic evidence of student learning and continual improvement. (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 473 073

# Combining the National Survey of Student Engagement with Student Portfolio Assessment

Juliette M. Stoering and Lina Lu

Institutional Research Analysts
Office of Institutional Research and Planning
Portland State University
P.O. Box 751
Portland Oregon 97207
Phone: 503-725-3432
Fax: 503-725-8755
e-mail: stoeringj@pdx.edu or lul@pdx.edu

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Vura

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)



Air Forum
Toronto, Ontario
Canada
June 2002

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

**Combining the National Survey of Student Engagement with Student Portfolio Assessment**

Colleges and universities grapple with how best to present evidence of authentic student learning to various audiences and stakeholders. They are also pressured to comply with third-party ranking systems based on data unrelated to learning outcomes. Two assessment activities have been particularly useful at one urban university with an innovative general education curriculum: the National Survey of Student Engagement (NSSE) and use of rubrics to score student portfolios. We analyzed the NSSE data in relation to the portfolio scores on four general education learning goals in an attempt to cross-validate self-report and authentic learning assessments. This project represented collaboration between faculty in the general education program and the institutional research office. Identifying the connections between diverse data sources and analyzing them in ways that respond to demands for accountability elevates the role of institutional researcher to that of a partner in the university's attempts to provide authentic evidence of student learning and continual improvement.

**Combining the National Survey of Student Engagement with Student Portfolio Assessment**

In response to demands for accountability, colleges and universities grapple with how best to present evidence of authentic student learning to various audiences and stakeholders. In addition, institutions of higher education are pressured to comply with controversial third-party ranking systems that perpetuate the myth that the true quality of an educational experience depends on institutional traits such as wealth, selectivity, and reputation. The data used in these ranking systems are largely unrelated to an institution's contribution to its students' learning outcomes (Graham & Thompson, 2001).

Dissatisfaction with the narrow view provided by third-party rankings and the need to provide more relevant information to the public has lead some schools to re-evaluate the institutional data they collect and report. Instead of counting students, faculty members, research grants, and alumni donations, these schools are attempting to measure student learning outcomes. This allows institutions to better communicate the contributions they make to student learning and target specific ways to improve the educational experience for their students.

Educators, administrators, and assessment experts have tried a variety of approaches to assessing student learning in order to bring student learning into the focus of the national conversation. These techniques include self-report items on national student surveys and evaluation of samples of student work.

Self-report surveys can provide cost-effective aggregate information about institutional effectiveness and, when an institution participates in a nationally distributed survey, can provide peer comparison information, as well. The value of these surveys depends on their relevance to specific institutional needs (Borden & Owens, 2001). Most measure students' opinions about their educational experiences, yet the true measure of institutional effectiveness is whether students are achieving learning outcomes (Angelo, 1999). In recent years, some national surveys have been developed to address more directly issues of student learning, but even these instruments are subject to questions about their validity (Borden & Owens, 2001).

Assessment using authentic materials, such as those collected in student portfolios, has gained in popularity over the last decade. Portfolio assessment is considered to be less biased and more holistic than traditional measures, such as tests and grades, and makes the expectations for the quality of student work explicit (Cook-Benjamin, 2001). In addition, portfolio assessment relies on multiple sources of information, which may improve its validity (Elbow, 1991); however, inter-rater agreement on portfolio scoring alone does not assure valid scoring (Heller, Sheingold, & Myford, 1998).

At one urban comprehensive university, the implementation of an innovative general education curriculum for undergraduates has made the need for good assessment practices increasingly important. One component of the general education curriculum is a yearlong, interdisciplinary, freshman program--Freshman Inquiry--that emphasizes building a foundation of communication skills for learning and expression and helps students learn and effectively use current information technologies. In the course, freshmen learn how the sciences, social sciences, humanities, and professional schools approach problems in different ways and how they work together to improve understanding of complex issues. Four learning goals are at the heart of the program: engaging in inquiry and critical thinking, using various forms of communication for learning and expression (with an emphasis on writing), gaining awareness of diversity, and appreciating ethics and social responsibility.

The culminating product of the first-year general education experience is a freshman portfolio. This portfolio is a collection of each student's work and reflections on their progress. Over the last two years, the faculty at the urban university have developed four rubrics corresponding to the general education learning goals and applied them to samples of freshman portfolios. Evidence of student learning fosters the credibility and acceptance of the program within the institution and communicates the high quality of the educational experience to the public.

Although many assessment activities are underway at this university, two are particularly useful in addressing the need for evidence of learning outcomes. These are the institution's

participation in the National Survey of Student Engagement (NSSE) and use of rubrics to score student portfolios. The NSSE data provides aggregate information on the extent to which student engage in behaviors related to positive educational outcomes (Kuh, 2001) and the portfolio assessment provides evidence of student learning in the classroom (Elbow, 1991; Cook-Benjamin, 2001). Separately, each source of information has proven useful to the institution, but given the national context, the university seeks to combine self-report with authentic assessment to provide a fuller picture of its contribution to student learning. Secondly, combining these data sources provides information about the validity of the evidence the university uses to demonstrate learning outcomes.

The Present Study and Its Limitations

The present study represented a first exploration of the relationship between NSSE data and data from rubric scoring of freshman portfolios. The general hypothesis was that rubric scores of students' portfolios would be correlated with student self-report on NSSE items related to the learning goals described in the rubrics. However, we were aware at the onset of the project of a number of limitations that could affect the likelihood of confirming the hypothesis. One limitation was that we had no prior research on which to base our prediction. Although it seemed intuitively likely that some items from the NSSE, given their face validity, would correlate with students' achievement of general education learning goals, both the NSSE and the PSU learning goal rubrics were fairly new instruments, and, to the best of our knowledge, previous research connecting NSSE results to authentic learning outcomes did not yet exist.

Secondly, the work of faculty in developing scoring rubrics for use with freshman portfolios and the work of the institutional research office in facilitating the NSSE administration were unconnected. Although general education faculty members were asked to encourage their students to respond to the NSSE and received a summary of the results, their familiarity with the instrument and investment in its use at the institution was minimal. Similarly, institutional researchers were only minimally aware of general education faculty's efforts to develop rubrics corresponding to the curricular learning goals. The rubrics were developed without considering

how they might correspond to the content or results of the NSSE. This is not to suggest that the NSSE items should have driven the process of developing the rubrics, but the research and reasoning behind the NSSE might have assisted the faculty in fine-tuning the rubrics and looking for commonalities between the two measures.

A third limitation was an issue of validity, and in effect, served as a broad research question of the study. Although items on the NSSE and the scoring rubrics appeared to share face validity, the question of whether students understand items on the NSSE and the scoring rubrics in the same way was an empirical question.

The final limitations of the study were related to practical matters. The process of identifying students who completed portfolios and responded to NSSE, obtaining their consent to review their materials, and completing the time-intensive task of scoring their portfolios reduced our sample size to the point that our analysis had limited the power to detect significant findings.

Merit of the Present Study

Despite these limitations, the project was worth pursuing for a variety of reasons. To the extent that the study could yield significant results, the findings would represent a contribution to institutional efforts to provide various constituents with a more complete understanding of student learning outcomes. Regardless of the specific findings of the study, any information about the connection between the NSSE and scoring rubrics would contribute to a better understanding of both measures. More specifically, the general education faculty could use such information to inform revisions of their rubrics.

Beyond the results, the study had merit for at least two other reasons. First, it modeled what should become a more common practice. Often, the decentralized nature of colleges and universities makes it difficult to keep abreast of the many assessment efforts and data collection efforts that take place. As a result, decentralized units do not pay enough attention to strategically coordinating their efforts so that they inform each other. The present study is an example of the role institutional research can play in identifying the connections between efforts and assist in strategic coordination. Second, the study represents collaboration between faculty

and institutional researchers. Identifying the connections between diverse data sources and analyzing them in ways that respond to needs for better information elevates the role of institutional researcher to that of a partner in efforts to provide authentic evidence of student learning. The benefits to the faculty and institutional researchers extend beyond the boundaries of the specific project.

## Method

### Sample

The sample consisted of 32 students who completed portfolios as a requirement of their freshman year general education program (i.e., Freshman Inquiry) and who responded to the NSSE during Spring 2001. A total of 80 freshmen had completed portfolios and responded to NSSE, but only students who could be contacted to provide written consent to review their materials were included in the sample.

### NSSE Data

Assessment experts developed the NSSE to assist higher education institutions in assessing the learning environment for their undergraduates. NSSE data indicate the extent to which students are engaged in good educational practices and what they gain from their college experience. Items on the NSSE ask about student behaviors that correlate with positive learning and personal development outcomes of attending university (Kuh, 2001).

The survey responds to the need for a new approach to assessing higher education that could lead to quality improvement. Educational effectiveness should be measured in terms of students' learning gains and such measurement should give institutions useful information about how to improve student learning. To date, NSSE data are the most comprehensive source of information that allows schools to benchmark their student learning outcomes and chart their progress in comparison to peer institutions (Graham & Thompson, 2001).

### Scoring Rubrics

Faculty who teach in the general education program developed scoring rubrics that correspond to the four main learning goals for students who complete Freshman Inquiry. The

learning goals are engaging in inquiry and critical thinking, using various forms of communication for learning and expression (with an emphasis on writing), gaining awareness of diversity, and appreciating ethics and social responsibility. Freshman Inquiry faculty have developed the four rubrics and applied them to freshman portfolios over the past two years. Each rubric employs either a four-point or a six-point scale. The writing and critical thinking rubrics use six points and the diversity of human experience and ethical issues and social responsibility rubrics use four points.

To date, results of the scoring rubrics have been limited to descriptive reports of aggregate scores. Analysis of the rubrics, including estimates of reliability and validity, do not exist, although plans are underway to conduct more formal testing of these measures.

Scoring Procedures and Data Preparation

Although Freshman Inquiry faculty had previously scored the portfolios of 10 students in the sample, 12 portfolios had not been previously scored. Three institutional researchers and two graduate assistants[1] underwent training with the Freshman Inquiry faculty to learn the rubric scoring system and to calibrate their scoring. This consisted of reviewing scoring documentation and then practicing the application of the rubrics during a one-hour session.

Two scorers reviewed each portfolio independently and their scores were averaged for a final score. In cases where the two scorers differed by more than a point in their rubric scores, a third reviewer independently applied the rubric to the portfolio and assigned a score. Of these three scores, the two closest scores were included in the average and the outlier was excluded. In cases where the third score was equidistant from the first and second score, all three were averaged together.

The Indiana University Center for Postsecondary Research and Planning compiled the NSSE data and prepared a data file ready for analysis to the urban comprehensive university, so data preparation consisted of adding the rubric scoring data to the NSSE file.

---

[1] *We would like to thank our colleague, David Burgess, and our graduate assistants, Lori Patterson and Andy Wang for their invaluable assistance in collecting and scoring the portfolios.*

## Results

We performed reliability testing on the scoring rubrics and produced descriptive statistics and correlations. Keeping in mind that our scoring procedure artificially inflated out alpha level, Table 1 shows that the reliability was acceptably high for three of the four rubrics. Critical thinking, Communication, and Diversity were fairly reliable measures, however, Ethics and Social Responsibility was less reliable. The descriptive statistics showed that scorers used the full range of the scale on each rubric, except for Communication. None of the portfolios received a score lower than '2' on the Communication rubric. Communication was significantly correlated with Critical Thinking ($r = .48$, $p < .01$) and Diversity ($r = .34$, $p < .05$), but not with Ethics and Social Responsibility ($r = .11$, $p > .05$).

Correlations were conducted between NSSE items and the four rubric scores. Overall, most of these were not significant at $p < .05$. Only two NSSE items shared modest correlations with Critical Thinking scores. The more frequently classmates worked together outside of class to prepare their assignments, the higher their Critical Thinking scores ($r = .36$, $p < .05$). In addition, Critical Thinking scores were somewhat related to student reports of how frequently they engage in various mental activities. Although we had hypothesized that critical thinking would be related to high-level mental activities (e.g., analyzing, synthesizing, evaluating, and applying theories), the statistics did not support this prediction ($p > .05$ for all correlations). However, the less engaged activity of rote memorization or repeating of facts or ideas in papers or on tests was negatively correlated with Critical Thinking ($r = -.39$, $p < .02$).

## Discussion

Higher education is searching for more authentic measures of student learning. We combined data from the NSSE with freshman portfolio scores to investigate whether the resulting information would have value at the program level and provide more complete evidence of learning outcomes. Although the learning goal rubrics applied to the portfolios and items on the NSSE appeared to have similar content, our results indicated little correlation between the two sets of data. Although we had hoped to find more, albeit weak, correlations, we were not

surprised by the results because of our understanding of the limitations of the methodology. Despite the lack of significant results, the outcomes of this research were valuable to our institutions and had implications for the role of institutional research.

Implications for Applying Rubrics to Freshman Portfolios

The analysis indicated that one of the rubrics, Ethics and Social Responsibility was not as reliable as the other rubrics. The Freshman Inquiry faculty is aware of this finding and plans to review and possibly modify the rubric. In addition, they plan to provide more training in its application to the portfolios.

Institutional Research and Freshman Inquiry faculty also discussed the correlations of the Communication rubric with Critical Thinking and Diversity of Human Experience. Because the evidence of critical thinking and appreciation of diversity was expressed through a written medium in the portfolios, the correlations were expected. If the reliability of Ethics and Social Responsibility were higher, we would have expected it to correlate with the Communication rubric as well. The correlations were relatively modest, accounting for 11.6% to 23.0% of the shared variance. This indicated that, although Critical Thinking and Diversity are related to Communication, for the most part, the three rubrics are measuring distinct abilities.

Implications for National Surveys

Freshman Inquiry faculty found the lack of correlation between the rubric scores and the NSSE data puzzling, despite their understanding of the limitations of the methodology. Two interpretations (in addition to the possibility that the small sample size reduced the power of the statistics to detect relationships) are under discussion at the institution. The first relates to the validity of each measure. Although faculty and institutional researchers recognize the overlap in content between the NSSE and the Freshman Inquiry learning goals, students might not. Students might interpret the meaning of items on the NSSE differently than they interpret the meaning of scoring rubrics used in Freshman Inquiry. The Freshman Inquiry faculty already makes an effort to word the rubrics in ways that students will understand them, and, in the future, will consider incorporating the wording of NSSE items in descriptions of the rubrics. They

speculate that the wording of the NSSE section that begins with "To what extent has your experience at this institution contributed to your knowledge, skills, and personal development in the following areas?" (and then includes items such as critical thinking and effective writing) does not cue students to reflect specifically on their experiences in Freshman Inquiry when responding. Freshman Inquiry faculty, with the help of institutional research, are considering conducting informal focus groups with students to better understand the reasoning they use in responding to the NSSE and whether students are aware of the items' connections to their Freshman Inquiry Experiences.

Questions about the validity of the NSSE for use in the general education program, leads to an ongoing debate about the utility of national surveys compared to institutionally developed surveys. One benefit of institutional surveys is that the institution has control over the wording of the survey and can test the validity of the items for use with its specific student population. Although much testing and effort goes into national surveys, they are constrained to generic language that can span a variety of institutions and students. In contrast, the availability of comparator data made possible by national surveys is a great advantage over institutional surveys. At our institution, this advantage outweighs the concerns over validity.

A second explanation of the lack of correlation between rubric scores and NSSE data is developmental in nature. Freshman portfolios demonstrate students' abilities to perform academic and intellectual exercises. The NSSE requires students to evaluate and report on their behaviors. At the freshman level, it is conceivable that fulfilling academic assignments is weakly linked to behavior. NSSE items cover behaviors and practice that are theoretically linked to positive learning outcomes, but it is unknown how long it takes for the practices to results in the desirable outcomes. After only three terms of exposure to the university experience, it may be too soon to expect that academic exercises demonstrated in portfolios would be closely tied to self-reported behavior or self-evaluation.

Faculty members in Freshman Inquiry and institutional research have discussed conducting a study of seniors' responses to the NSSE and their work in Capstone, a culminating

general education experience. The prediction would be that NSSE data and achievement of general education learning goals would be more closely related for seniors than for freshmen. However, the final products in Capstone are group projects, so it would be difficult to identify individual accomplishment of the learning goals. At present, the faculty is working with institutional research to consider the feasibility of collecting and applying rubrics to other work samples of seniors in Capstone classes.

Implications for Institutional Research

Combining the data from the NSSE with portfolio scores represents collaboration between faculty in the general education program and the institutional research office. Discussion about this project has increased awareness and understanding about both the rubric scoring and NSSE data. Furthermore, it fostered greater awareness of the resources available through the institutional research office.

The Freshman Inquiry faculty has come to recognize the helpfulness of including institutional researchers as consultants to their assessment efforts. They also recognize that relying on institutional research for advice about appropriate and rigorous methodologies increases the credibility of assessment results they disseminate to the campus community. This is an example of what Matier, Sidle, and Hurst (1995) described as the changing role of institutional researchers. Increasingly more often, our office provides substantive analysis that reaches beyond the traditional activities of data collection and reporting. By identifying the connections between diverse data sources and analyzing these data in ways that respond to demands for accountability, we elevate the role of institutional researcher to that of a partner in the university's attempts to provide authentic evidence of student learning and continual improvement.

**References**

Angelo, T. A. (1999). Doing assessment as is learning matters most. AAHE Bulletin, 51(9), 3-6.

Borden, V. M. H., & J. L. Z. Owens. (2001). Measuring quality: Choosing among surveys and other assessments of college quality. Washington, DC: American Council on Education; Tallahassee, FL: Association for Institutional Research.

Cook-Benjamin, L. (2001). Portfolio assessment: Benefits, issues of implementation, and reflections on its use. Assessment Update, 13(4), 6-7.

Elbow, P. (1991). Foreword. In P. Belanoff & M. Dickson (Eds.), Portfolios: Process and Product (pp. ix-xxiv). Portsmouth, NH: Boynton/Cook.

Graham, A., & Thompson, N. (2001, September). U.S. News' college rankings measure everything but what matters. And most universities do not seem to mind. The Washington Monthly [On-line serial]. Available url: http://www.washingtonmonthly.com/features/2001/0109.graham.thompson.html.

Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the National Survey of Student Engagement. Change, 33(3), 10-17, 66.

Matier, M. W., Sidel, C. C., & Hurst, P. J. (1995). Institutional researchers' roles in the twenty-first century. In P. T. Terenzini (Series Ed.) & T. R. Sanford (Vol. ed.), New directions for institutional research: Vol. 85. Preparing for the information needs of the twenty-first century (pp. 75-84). San Francisco: Jossey-Bass.

Table 1

Reliability and Descriptive Statistics of Rubric Scores

| Rubric | Score Range | Score 1 | | Score 2 | | Final Score | | Alpha | Standardized Alpha |
|---|---|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD | | |
| Inquiry & Critical Thinking | 1 – 6 | 2.91 | 1.35 | 2.91 | 1.40 | 2.91 | 1.29 | .86 | .86 |
| Communication | 1 – 6 | 3.75 | 0.80 | 3.50 | 1.02 | 3.62 | 0.82 | .76 | .77 |
| Diversity of Human Experience | 1 – 4 | 2.56 | .80 | 2.56 | .95 | 2.56 | 0.78 | .74 | .74 |
| Ethical Issues & Social Responsibility | 1 – 4 | 2.12 | .83 | 1.97 | .82 | 2.03 | 0.67 | .50 | .50 |

**U.S. Department of Education**
*Office of Educational Research and Improvement (OERI)*
*National Library of Education (NLE)*
*Educational Resources Information Center (ERIC)*

# NOTICE

## Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").